

## NCDC Mass Storage Systems and Technologies

**Dick Davis**

NOAA/National Climatic Data Center

Federal Bldg

Asheville, NC 28801

Phone: (704) 271-4384

FAX: (704) 271-4246

ddavis@NCDC.noaa.gov

### 1. Size Of Holdings:

Current holdings at NCDC are 107.8 terabytes of digital data and about 0.3 terabytes of manuscript data.

The nexrad radar system is expected to generate approximately 88 terabytes per year by 1996.

### 2. Nature Of Holdings:

Basically environmental data. Climatological observations at varying temporal scale from 1 minute to monthly values for both surface and upper air. Analyzed grids of surface and upper air, summarized climatological information, surface marine observations and gridded values, selected satellite data. There are approximately 400 different tape decks within the archives.

### 3. How Long Site In Existence

NCDC was first established in 1938 as a wpa project to use punched cards to tabulate climatological information. In 1952 the center was relocated to asheville, nc. So, for 14 years we were the new orleans tabulation unit, then for the next 41 years we became known as the national weather records center, the national climate center and then in 1976, the national climatic data center.

### 4. Popularity Of Data Sets:

The popularity of our basic climatological data sets, hourly surface, summary of the day, and hourly precipitation has remained essentially constant. The top ten are:

- Surface airways hourly
- Daily cooperative summary of the day
- First order summary of the day
- Surface/land summary of the month
- Datsave surface hourly
- NCDC us upper air
- Hourly precipitation data
- National solar radiation data base
- Mixing height studies
- Surface marine observations

## **5. Media/Technology Used For Storage:**

Current policy is that archives will be on 3480 square cartridges. We have just completed a three year effort to copy all round tapes in our primary archives to the 3480 media. This has included merging two or three tapes to one cartridge where possible.

In addition to the reduced storage space needed, the read/write reliability is much enhanced. With increasing acquisition of archive files, we have been able to maintain a viable tape library without an increase in storage area.

We are also involved in a data rescue effort, transferring satellite data from round tapes to cartridge tapes. Over 50,000 tapes have been "rescued" so far and the remaining 20,000 or so tapes will be processed prior to July 1994.

The choice of storage media has influenced the distribution of data but generally user capabilities have also kept pace with the newer tape technologies. While we no longer provide data on 7-track tape, we do maintain the ability to furnish data on 9-track tape at 1600 or 6250 bpi, as well as on 3480 cartridges.

We also are able to extract data from the archives and provide data to customers on 8 mm tape or 3 1/2" and 5 1/4" diskette. These are the preferred media by many customers who have pcs and who do not work with very large data files.

## **6. Volume Distributed Per Month:**

447 media units/month over the past two years.

## **7. Mode Of Distribution:**

Principal modes of distribution remain as magnetic tape, diskette, 8 mm tape, or cd-rom by mail. On-line capability is increasing and is available through internet. Most of these on-line data sets are special projects such as profiler, or the most recent (perhaps month) period of record from the principal climatological files e.g. Summary of the day or hourly airways observations.

The on-line system also includes several inventory sets and in some cases allows the user to request copies of data sets to be copied off-line, and mailed to him/her.

Another increasingly popular mode of dissemination is by spectra-fax. Several of the most requested publications and data sets are kept on line and the user has only to dial into the fax machine, enter his account number or credit card and then specify the data he wants. The hard disc is searched and the data transferred directly. No human intervention at the NCDC is required.

## **8. Most Frequently Encountered Problems:**

**Archives:** acquisition of "stranger" tapes that do not conform to stated formats, have internal labels that conflict with our tape management system, or that do not contain the data purported to be on the tape.

"older" tapes that are difficult or impossible to read on the high speed drives of the main frame and which must be copied on other, slower speed drives, before converting to cartridge tapes.

Binary tapes with ebcdic labels. These must be copied onto other round tapes in order to convert the labels to ascii and then copying to cartridge tapes can proceed.

Most problems encountered in using the 8 mm tapes from the nexrad system seem to be procedural or system errors rather than problems with the tapes or drives themselves. We have, however, encountered some difficulties during the write process that have been attributed to tape debris. This whole system is in its infancy and it remains to be seen how viable the 8 mm technology is for continuous drive operation and long term retention. Economics dictate the use of this technology.

Customers: we guarantee readability of digital data for a period of 60 days. Very few customers experience difficulty with the output media. Probably the most frequent complaint is with the documentation that is provided.

## **9. Type Of Media Requested/Used:**

Over the past two years requests for floppy disks have over taken those for magnetic tape.

Requests for 1600 bpi tapes have virtually disappeared. Most tape customers still want round reels at 6250 bpi although we do have some who request 3480 cartridges.

There has also been an increase in the number of customers asking for data on 8 mm tapes.

We will soon have seven cd-roms available. There has been a 900% increase in requests for data on this media over the past two years.

## **10. Evolution Of Media:**

Through the years, storage media has basically kept pace with newest technology. This has provided the opportunity to systematically migrate data sets in order to ensure the readability of the data, as well as decrease the number of media units required to hold the ever increasing amount of data in the archives. The progression has been:

- Punched cards
- 7-track 3/4" mag tape
- 7-track 1/2" mag tape at 200 bpi
- 7-track 1/2" mag tape at 556 bpi
- 7-track 1/2" mag tape at 800 bpi
- 9-track 1/2" mag tape at 800 bpi
- 9-track 1/2" mag tape at 1600 bpi
- 9-track 1/2" mag tape at 6250 bpi
- 3480 cartridge mag tape
- 8 mm helical scan tape\*

\* this media is being used for archiving of nexrad data only.

We are also producing and distributing special data sets on cd-rom.

## **11. Wish List:**

We are looking at an hierarchical mass store subsystem that can be upgraded as needed and as funds permit. This piecemeal approach, may not be ideal, but in the real world sometimes one has to use innovative techniques in order to secure to the desired end result.

A dual or quadruple creo optical tape system that could be used both as permanent storage for the archives and as an on-line access for our principal data sets, using raid technology.

Development of an optical tape system using 1/2" film cartridges that could be used in a robotic system. This would be more salable than the 35 mm format now in existence. The recording densities and access times developed by the creo corporation, coupled with the shelf life of the media make this a most attractive approach.

And of course, my pet wish - a truly operational holographic storage/recall system. To an archivist, this would be the ultimate permanent media providing high density recording and nearly indestructible data files.

## **12. Words Of Wisdom:**

Caution but not inertia as you attempt to solve the problems of storing multi-terabytes of data. The next technological break-through is *always* just around the corner. At some point you have to go with your best intuition and declare that you have reached the corner.

Changing technology must be kept in mind however, and one should assume the mantle of omniscience, planning for the inevitable migration to yet another "new" system.